School of Mathematics and Statistics,
University of Sydney, Australia
georg.gottwald@sydney.edu.au
Georg Gottwald

June 25, 2026

Prof. Christian Franzke
Editor, NPG

RE: egusphere-2025-6535: Contrasting different noise models for representing westerly wind bursts in a recharge oscillator model of ENSO

Dear Editor,

Thank you for the opportunity to resubmit our work. We would like to thank the referees for their constructive criticism. We revised the paper throughout. Please find below our response to the referee reports.

Best wishes,

Georg Gottwald, Eli Tziperman and
Alexey Fedorov

Response to Reviewer 1

Thank you very much for your careful reading and the detailed and valuable comments.

This manuscript describes the role of various noise approximations to the role of Westerly Wind Bursts (WWBs) on the development of large amplitude El Nino events. The finding that conditional additive and multiplicative (CAM) noise is a reasonable approximation to intermittent WWBs with impacts on asymmetries between El Nino and La Nina events is perhaps not unsurprising given the literature on applications of noise in linear inverse models (LIMs). However the demonstration within a simplified framework of the recharge oscillator model is somewhat novel. In general, this work is well described, clear and accessible to a wide audience. It is entirely appropriate for NPG.

We thank the referee for the positive assessment.

Given the effort to produce the conditional noise model (CON) I am somewhat confused that this is not at all mentioned in the abstract nor why the paper only has a focus on large amplitude El Nino and not on the ability to reproduce the observed La Nina amplitudes.

We have expanded the abstract and further highlighted that our CON model better represents certain dynamical signatures associated with WWBs:

> To cover the full spectrum of warm events, we propose a conditional noise model in which the wind stress is modelled by additive Gaussian noise for sufficiently small SSTs and by additive CAM noise once the SST exceeds a certain threshold. We show that this conditional noise model captures observed bulk statistical properties of ENSO equally well as the commonly used multiplicative Gaussian red noise model, but additionally better reproduces dynamical signatures such as the increased number of WWBs preceding large El Niño events.

Specific comments:

Paragraph 1; Line 1: Reference [1] is to an ECMWF tech. report from 1997. I suggest replacing this with a more up to date reference that takes into account the most recent decades.

Thanks for this catch. we now cite instead Harrison and Vecchi (1997), Lengaigne et al. (2004), and Yu and Fedorov (2022), with further recent literature in the following parts of the introduction.

Paragraph 2; Line 1: What is the "inferred measure of the time-integrated wind stress associated with WWBs"? Is it simply wind stress anomalies? Please explicitly state what is calculated in Figure 1. Also why use CESM rather than a reanalysis e.g. ERA5. I would further suggest combining the observed time integrated wind stress due to WWBs with the observed NINO3 index and demonstrate i.e. quantify the relationship.

The information was buried in the caption of Figure 1. We now define the measure in the main text as well.

> To motivate our proposed noise model, consider Figure 1a, which shows an inferred measure of the time-integrated wind stress associated with WWBs from a numerical simulation of a global configuration of the Community Atmospheric Model (Conley et al., 2010). ... Specifically, we show $\Delta t_{\mathrm{WWB}}\, U_{\mathrm{WWB}}^2$, where Δt_{WWB} is the WWB duration and U_{WWB} is the surface wind speed strength. This measure attempts to represent the effect of these events on the ocean, which depends on both their amplitude and their duration.

We admittedly don't have access to the equivalent information on the times and magnitude of observed WWB events, but have now added a reference indicating that the modelled events are sufficiently realistic for the purpose of this illustrative figure:

> Lian et al., 2018 show that even an earlier version of the Community Atmospheric Model simulates WWBs quite realistically.

Page 2: regarding Gaussian multiplicative noise on East Pacific SST, please provide references to the "... emerging consensus in the ENSO literature".

We added a reference to a recent review article by Vialard et al., 2025 and to Perez et al., 2005.

Page 3: Last sentence before section"Noise Models": missing section number. i.e., "in section ? we will discuss ..."

This should have been Section 3 and is now corrected. Thank you.

Sentence following Equation 8: ν_2 is undefined.

This should have been ν_1 and is now corrected. Thank you for noting this.

Page 6: Combine Fig3 with new Fig1 observed WWBs.

Done.

Fig 4: Please quantify the differences between observed and estimated variance and skewness. By eye, it appears that OU is a better fit than either CAM or CON.

We now include several statistical tests (two-sample Kolmogorov-Smirnov test, RMS error and Wasserstein distance) to quantify the differences between the three noise models (OU, CAM, and CON). It turns out that OU and CON are comparable, and CAM is performing slightly less (maybe not unexpected as by design it generates extreme events). We now explain,

> Figure 6 shows the corresponding histograms for the SST T_e and the power spectrum together with the corresponding curves for the observed NINO3 index. It is seen that all three noise models are capable of reproducing the global observed ENSO statistics reasonably well. To quantify to what degree each noise model can reproduce these statistics, we have compared several diagnostics. We performed a two-sample Kolmogorov-Smirnov test to check if any of the RO models generates SST data that could have been drawn from the distribution implied by the observational NINO3 index data. The null hypothesis was rejected for all three noise models, with CAM performing worst, as to be expected, as CAM noise is designed to generate only extreme warming events, rather than a wide range of events. The p-value of the OU model and the CON model are comparable with $p = 0.0114$ and $p = 0.0103$, respectively. We further estimated the Wasserstein distance between the synthetic T_e model output and the NINO3.4 index, which measures the difference between their respective distributions. The OU model and the CON model again exhibit similar Wasserstein distances from observations, of 0.0552 and 0.0569, respectively. The data generated by the CAM noise exhibits again a larger Wasserstein distance from observations, of 0.1234, which is to be expected as these data are heavy tailed with several extreme events. To quantify the differences in the generated power spectra, we used a simple mean-square error. Here all three models show roughly the same error with 0.1344, 0.1091 and 0.115 for OU, CAM and CON, respectively. We remark that we did not optimize the equation and noise parameters to produce optimal fits with the observations but rather used a visual assessment.

Fig5: I do not see the relevance of these timeseries segments. Please provide a quantitative measure or drop the figure.

We added quantitative measures as described in the response to the previous question, but felt it would be best to keep the time series to allow the reader a direct visual impression of the differences (or lack thereof) between the noise models. We also discuss the time series more explicitly now:

> Typical time series of the East Equatorial Pacific SST, T_e, are shown in Figure 4. We show some statistical characterizations of these time series below. Yet one might argue just based on these snapshots that multiplicative OU (Figure 2b) and the CON noise (Figure 2d) seem to produce similar and reasonable ENSO simulations, as also supported by the statistical analyses below.

Fig8: Please provide a table with the dates of the chosen events. Further, it is quite obvious that many of these events are not monotonically increasing until at least t=-6 months. Choosing only the 4 most extreme cases does not strengthen the argument when all chosen cases are exceeding the threshold.

Thank you for this helpful comment. Given this feedback, we reconsidered and decided to remove the discussion on the monotonicity measure. This is also a result of the fact that the monotonicity measure could not be used to favor or disfavor any of the three noise models when comparing to the observations.

Response to Reviewer 2

Thank you very much for your careful reading and the detailed and valuable comments.

Major Comments:
The Introduction is currently somewhat unclear. (1) The authors need to explain the limitations of the classical multiplicative Gaussian-noise approach more explicitly. For example, is the main motivation that an OU process cannot adequately represent the highly nonlinear and intermittent nature of westerly wind bursts (WWBs)?

This is correct. Our aim is to show that while Gaussian noise and CON noise are equally well suited to reproduce global statistical features of ENSO, CON provides a better representation of WWBs. We now contrast in the new Figure 3 the noise signals of our CON model with the multiplicative noise forcing employed by Vialard et al 92025) as an example of the current state of the art. This figure clearly shows that Gaussian noise (even when multiplicative) is very different from the WWB time series in Fig. 1 and is therefore not appropriate when one is interested in reproducing sporadic peaky WWB events. We now explain more explicitly in the introduction,

> In this work, we examine an alternative to this stochastic forcing formulation, which provides an improved representation of the effects of WWBs on ENSO. We show that while Gaussian noise reproduces the global statistical features of ENSO, an alternative formulation is required to better represent the character of WWBs.

(2) The transition from CAM noise to the proposed CON model feels a bit abrupt. The authors state that CAM noise is "more appropriate" for warm environments, but this claim would benefit from clearer literature support.

We now make this argument more explicit in the introduction and support it with a few papers about both observational analysis and related modelling results:

> We wish to make the case that CAM noise is a more appropriate noise model to model WWBs that are known to occur with a developing El Niño warming, further amplifying large El Niño events (Yu, Weller, and Liu, 2003; Tziperman and Yu, 2007; Liang and Fedorov, 2021). Stochastic wind forcing over cooler oceanic environments may be more appropriately captured by Gaussian noise.

The manuscript needs to better discuss the limitations of using either a purely OU/Gaussian noise model or a multiplicative CAM-noise model. For instance, is the CON model introduced because CAM noise alone produces unrealistic La Niña behavior or cold-phase statistics? Without showing where the simpler OU-only or CAM-only approaches fail, the propose for the CON model may appear somewhat ad hoc, mainly introduced to fit the NINO3 PDF rather than as a physically motivated improvement. I suggest that the authors clarify this logic in the Introduction and, if possible, include a brief comparison showing the limitations of the OU-only and/or CAM-only approaches.

Thank you for the suggestion, which we have addressed in several ways. First, in the introduction, we now explain,

> In this work, we examine an alternative to this stochastic forcing formulation, which provides an improved representation of the effects of WWBs on ENSO. We show that while Gaussian noise reproduces the global statistical features of ENSO, an alternative formulation is required to better represent the character of WWBs.

That is, the issue is not that OU does not produce satisfactory El Niño or La Niña events, but that it does not have the character of WWBs.

We further address the issue in the new Figure 3 which shows the stochastic noise for CON and OU and demonstrates the similarity of CON to the WWB time series in Figure 1.

Since the authors argue that WWBs resemble CAM noise, it would also be helpful to clarify whether this type of noise has been applied to ENSO models in previous studies, and to strengthen the physical or literature-based justification for using CAM/CON noise in the RO model instead of more traditional noise schemes.

CAM noise has indeed be used in the context of ENSO which we were not aware of at our first submission. Martinez-Villalobos et al (2019) have shown that a multivariate model driven by CAM can be used to account for the observed asymmetry in ENSO. We now cite their work.

For Figures 4 to 6, the Results section lacks clarity regarding the advantages of the proposed models. In Figure 4, the traditional OU process (4a) appears to have a smaller variance, and its simulated mean (black dot) is closer to the observations (red dot) than the CON model (4c). In Figure 6, based on the PDF results, the OU process seems more consistent with observations, as the simulated curve stays mostly within the 5-95% confidence interval (blue and beige bars) if I understand correctly. Conversely, CAM and CON models show significant deviations at the peak of the PDF.

The revised manuscript is careful to explain that the advantage of the proposed WWB model is not an improved ENSO simulation in the recharge oscillator model in terms of reproducing global statistical features. Rather, both models do fine, except that the noise produced by the CON model is much more similar to actual WWBs. We make the case that it is important to get the noise right, not just the global statistical features of ENSO, as well as to recover certain dynamical features of WWB-induced large El Niño events such as the increased number of WWBs preceding a large event.

Beyond that, we now include several statistical tests (two-sample Kolmogorov-Smirnov test, RMS error and Wasserstein distance) to quantify the differences between the three noise models (OU, CAM, and CON). It turns out that OU and CON are comparable, and CAM is performing slightly less (maybe not unexpected as by design it generates extreme events). We now explain,

Figure 6 shows the corresponding histograms for the SST T_e and the power spectrum together with the corresponding curves for the observed NINO3 index. It is seen that all three noise models are capable of reproducing the global observed ENSO statistics reasonably well. To quantify to what degree each noise model can reproduce these statistics, we have compared several diagnostics. We performed a two-sample Kolmogorov-Smirnov test to check if any of the RO models generates SST data that could have been drawn from the distribution implied by the observational NINO3 index data. The null hypothesis was rejected for all three noise models, with CAM performing worst, as to be expected, as CAM noise is designed to generate only extreme warming events, rather than a wide range of events. The p-value of the OU model and the CON model are comparable with $p = 0.0114$ and $p = 0.0103$, respectively. We further estimated the Wasserstein distance between the synthetic T_e model output and the NINO3.4 index, which measures the difference between their respective distributions. The OU model and the CON model again exhibit similar Wasserstein distances from observations, of 0.0552 and 0.0569, respectively. The data generated by the CAM noise exhibits again a larger Wasserstein distance from observations, of 0.1234, which is to be expected as these data are heavy tailed with several extreme events. To quantify the differences in the generated power spectra, we used a simple mean-square error. Here all three models show roughly the same error with 0.1344, 0.1091 and 0.115 for OU, CAM and CON, respectively. We remark that we did not optimize the equation and noise parameters to produce optimal fits with the observations but rather used a visual assessment.

The authors claim on page 6 that "... all three noise models are capable of reproducing the global statistics reasonably well," but this is a qualitative statement. I suggest using a quantitative metric (e.g., Kolmogorov-Smirnov test or RMSE) between observations and models to justify this claim.

Thank you for this fair statement. As explained above, we have now included several test statistics: to compare the empirical histograms we employed a two-sample Kolmogorov-Smirnov test, an RMS distance in the power spectrum and calculated the Wasserstein distance between the observed NINO3 index and our synthetic T_e data. Please see above quote from the paper in response to the previous question, with the details.

If the OU model provides a better statistical fit, the authors should explicitly discuss the trade-off between "statistical fit" and "dynamical realism."

This is exactly our aim. Thank you for putting it so succinctly which we failed to do in the previous version. We now highlight this aspect at several places. For example, we now write in the Conclusion,

> All three models were able to explain the spectrum, the global ENSO statistics such as the empirical histogram, the variance, and the skewness of the observed NINO3 time series. The various noise models, however, give rise to different noise characteristics that are not all appropriate for representing WWBs. The point we are making here is that, from an understanding point of view, it is not sufficient for the noise model to produce ENSO's characteristics, but the noise itself should resemble the observed WWB characteristics. Furthermore, characteristics such as having more WWBs preceding large El Niño events should be accounted for as well.

A recurring issue is the absence of direct observational benchmarks in the dynamical analysis in Figures 5, 7, & 9. Figure 5: The simulated SST time series is shown in isolation. Without a side-by-side comparison with the observed NINO3 time series, it is impossible to evaluate if the "morphology" or "intermittency" has been improved.

Thanks for this helpful comment. We have added the last 60 years of the observed NINO3 index to Fig. 4 (previously Figure 5) in addition to the fuller extent shown in Figure 1. We have deleted the monotonicity measures that was shown in Figure 9, and we have added quantitative measure of the distance from observations as mentioned above, including two-sample Kolmogorov-Smirnov test, the Wasserstein distance between the synthetic T_e model output and the NINO3 index, and a mean-square error estimate.

Figure 7: I think this is one of the most interesting results of the paper, yet it lacks an observational equivalent. The authors should calculate the number of strong WWBs preceding historical El Niño events and include this as a reference histogram.

We agree that this is an interesting result. We did not add the observed equivalence as that would require a quantification of observed WWBs that is non trivial. We did add some references on the hopefully well-known phenomenon of a positive feedback between the SST and WWBs, including Tziperman and Yu (2007) and Gebbie et al. (2007) and others.

Figure 9: The monotonicity measure μ is an innovative diagnostic. However, to validate its utility, the authors need to calculate μ from the observed NINO3 index (e.g., for the 22 events mentioned) and mark these values on the plots. This would provide a definitive test of whether CAM/CON noise truly captures the "fingerprint" of real-world events better than the OU process.

Thank you for suggesting these experiments which we now conducted. Unfortunately, given the short observed record, we felt we could not make a strong case that this measure points to any noise model as being able to better reproduce the observed signatures. We have therefore removed the related figures and discussion.

Minor Comments:

Figure 6 caption: The caption mentions a "green line," but I only see blue bars and red/black curves. I assume "green" refers to the "blue"? Please ensure the colors in the caption match the figure. Additionally, please define exactly what the red line represents in all subplots of the figure.

Yes, this was a typo. It should have been “blue”. Corrected now, thank you!

Missing Line Numbers: The manuscript lacks line numbers. Please include them in the revised version.

Done.

Parameter Sensitivity: The parameters in Table I appear highly tuned for each model. Please discuss whether the improved monotonicity in the CON model is a robust feature of CAM noise or sensitive to specific parameter selections.

The parameters were chosen to *roughly* fit the bulk statistics of the observed ENSO SST signal, i.e., the empirical histogram, power spectrum and variance-skewness histogram. They were not tuned for the monotonicity test (but this measure has now been removed from the manuscript). In Figure 8 (previously Figure 10) we used a CON model without the asymmetric thermocline response, with $\gamma_+ = \gamma_-$ showing the same qualitative behaviour illustrating the robustness of the qualitative features.

References

Conley, A.J., R. Garcia, D. Kinnison, J.F. Lamarque, D. Marsh, M. Mills, A.K. Smith, S. Tilmes, F. Vitt, H. Morrison, Philip Cameron-Smith, William D. Collins, Michael J. Iacono, Richard C. Easter, Steven J. Ghan, Xiaohong Liu, Philip J. Rasch, and Mark A. Taylor (2010). *Description of the NCAR Community Atmosphere Model (CAM 5.0)*. Tech. rep. National Center for Atmospheric Research.

Gebbie, G., I. Eisenman, A. T. Wittenberg, and E. Tziperman (2007). "Modulation of westerly wind bursts by sea surface temperature: A semi-stochastic feedback for ENSO". In: *Journal of the Atmospheric Sciences* 64, pp. 3281–3295. DOI: 10.1175/JAS4029.1.

Harrison, D. E. and G. A. Vecchi (Dec. 1997). "Westerly wind events in the tropical Pacific, 1986-95". In: *Journal of Climate* 10.12, pp. 3131–3156.

Lengaigne, Matthieu, Jean-Philippe Boulanger, Pascal Delecluse, Christophe Menkes, Eric Guilyardi, and Julia Slingo (2004). "Westerly wind events in the tropical Pacific and their influence on the coupled ocean–atmosphere system: A review". In: *Earth Climate: The Ocean–Atmosphere Interaction*. Ed. by Chunzai Wang, Shang-Ping Xie, and James A. Carton. Vol. 147. AGU Geophysical Monograph Series. Washington, DC: American Geophysical Union, pp. 49–69. DOI: 10.1029/147GM03.

Lian, Tao, Youmin Tang, Lei Zhou, Siraj Ul Islam, Chan Zhang, Xiaojing Li, and Zheng Ling (2018). "Westerly wind bursts simulated in CAM4 and CCSM4". In: *Climate Dynamics* 50.3-4, pp. 1353–1371.

Liang, Yu and Alexey V. Fedorov (2021). "Linking the Madden–Julian Oscillation, Tropical Cyclones and Westerly Wind Bursts as Part of El Niño Development". In: *Climate Dynamics* 57.3–4, pp. 1039–1060. DOI: 10.1007/s00382-021-05757-1.

Perez, C. L., A. M. Moore, J. Zavala-garay, and R. Kleeman (Dec. 2005). "A comparison of the influence of additive and multiplicative stochastic forcing on a coupled model of ENSO". In: *Journal of Climate* 18.23, pp. 5066–5085.

Tziperman, E. and L. Yu (2007). "Quantifying the dependence of westerly wind bursts on the large scale equatorial Pacific SST". In: *Journal of Climate* 20.12, pp. 2760–2768. DOI: 10.1175/JCLI4138a.1.

Vialard, Jérôme, F-F Jin, Michael J Mcphaden, Alexey Fedorov, Wenju Cai, S-I An, D Dommenget, X Fang, MF Stuecker, C Wang, A. Wittenberg, S. Zhao, F. Liu, S.-K. Kim, Y. Planton, T. Geng, M. Lengaigne, A. Capotondi, N. Chen, L. Geng, S. Hu, T. Izumo, J.-S. Kug, J.-J. Luo, S. McGregor, B. Pagli, P. Priya, S. Stevenson, and S. Thual (2025). "The El Niño Southern Oscillation (ENSO) recharge oscillator conceptual model: Achievements and future prospects". In: *Reviews of Geophysics* 63.1, e2024RG000843.

Yu, L., R. A. Weller, and T. W. Liu (2003). “Case analysis of a role of ENSO in regulating the generation of westerly wind bursts in the Western Equatorial Pacific”. In: *Journal of Geophysical Research* 108.C4, 10.1029/2002JC001498.

Yu, Sungduk and Alexey V. Fedorov (2022). “The essential role of westerly wind bursts in ENSO dynamics and extreme events quantified in model “wind stress shaving” experiments”. In: *Journal of Climate* 35.22, pp. 7519–7538. DOI: 10.1175/JCLI-D-21-0401.1.